A contrastive look at the processing of metaphor and hyperbole *

Felicity Deamer, Nausicaa Pouscoulous and Richard Breheny

Abstract

Lexical pragmatic accounts of figurative language interpretation (Carston 2002, Sperber and Wilson 2006, Wilson 2003, Wilson and Carston 2007, Recanati 1995, 2001, 2004) propose that the interpretation of metaphor and hyperbole involves constructing an ad hoc concept on the basis of the concept encoded by the metaphor or hyperbole vehicle. This paper discusses these accounts and investigates their main empirical prediction: that the interpretation of both metaphor and hyperbole involves the construction of an ad hoc concept by enhancing properties of the metaphor or hyperbole vehicle that are relevant for interpretation, while suppressing those that are irrelevant. We carried out a reading-time study contrasting metaphor and hyperbole in order to investigate this claim. Participants read a series of target sentences preceded by a metaphor, a hyperbole, or a literal priming context sentence. For half the participants, the context sentences were short (5-13 words), and for the other half they were longer (55-86 words). Participants took significantly longer to read target sentences in the long context condition than in the short context condition and reading times of target sentences were faster when preceded by a hyperbole or literal biasing context sentence than when preceded by a metaphor biasing context sentence.

Keywords: Metaphor, Hyperbole, language comprehension, suppression, processing

1 Introduction

Traditionally, metaphorical interpretations have been considered more complex than literal interpretations. Based on a Gricean approach to meaning (Grice 1989), a metaphorical interpretation was considered to be derived from the literal interpretation of a sentence, which must be arrived at first (Ortony 1979, Gentner and Wolff 1997). Currently, certain theorists are pursuing a new direction of thinking with regards to metaphor and figurative language processing in general (see, e.g., Carston 2002, Sperber and Wilson 2006, Wilson 2003, Wilson and Carston 2007, Recanati 1995, 2001, 2004). One of the premises of this new approach is that there is no default literal interpretation which must be arrived at before a more relevant figurative interpretation can be derived. They propose that the comprehension of metaphor involves the on-line use of existing stable concepts in the construction of ad hoc, occasion specific concepts which denote the broader metaphorical meaning. They argue that although the lexically encoded stable concepts are accessed, the literal interpretation of the whole utterance does not need to be derived. The Relevance Theoretic (RT) unified account (Carston 2002, Sperber and Wilson 2006, Wilson 2003, Wilson and Carston 2007) and Recanati's (1995, 2001, 2004) account, both propose that hyperbole and metaphor are the result of a process of on-line ad hoc concept construction, and that the resulting broadened interpretations contribute to the proposition expressed by the utterance, rather than being conveyed implicitly.

Given its importance for any pragmatic theory, metaphor raises a host of interesting issues, both theoretical and psychological. Is metaphor part of the proposition expressed? Must a metaphoric meaning of a word always be derived from a literally encoded meaning?

Is the on-line derivation of metaphor accurately captured by the processes and mechanisms described in pragmatic theories?

In recent decades many studies have investigated the psychology of metaphor interpretation (see, among many others, Ortony, Schallert, Reynolds and Antos 1978, Inhoff, Lima and Carroll 1984, Blasko and Connine 1993, Glucksberg, Gildea & Bookin 1982, McElree and Nordlie 1999, Rubio-Fernandez 2004, 2007, 2008). Most of which have challenged the Gricean-inspired paradigm in which metaphors are derived via literal meaning. In comparison to metaphor, other types of figurative language have received little empirical attention.

The recent proposals by Relevance theorists and Recanati are for a unified lexical pragmatic account of figurative language, which falls in line with the findings of the studies on metaphor mentioned above (Ortony et al. 1978, Inhoff et al. 1984, Blasko and Connine 1993, Glucksberg et al. 1982, McElree and Nordlie 1999, Rubio-Fernandez 2004, 2007, 2008), and which encompasses all tropes in which the lexically encoded meaning is broadened such as metaphor, hyperbole, approximation, and category extension, as well as narrowing. These proposals provide strong motivation to further investigate these linguistic phenomena in order to address several questions which remain unanswered: To what degree are the processes involved in the comprehension of these different tropes the same? Are the precise cognitive mechanisms involved the same for all tropes? Are all figurative interpretations as rich? Do interpretations of some tropes require more attentional resources than others?

Lexical priming studies (Rubio-Fernandez 2004, 2007, 2008) have provided some insight into the cognitive processes underpinning metaphor comprehension; Rubio-Fernandez' findings indicate that metaphor irrelevant superordinates lose their activation between 400 and 1000ms as a result of active suppression. Similarly, sentence verification studies (Gernsbacher, Keysar, Robertson, and Werner 2001), also indicate that metaphor irrelevant features of the metaphor vehicle are suppressed and metaphor relevant superordinates are enhanced during processing.

Studies investigating whether similar suppression and enhancement effects can be observed during the processing of other types of broadening, such as hyperboles and approximation, have not yet been carried out. A study contrasting the interpretation processes involved in metaphor and hyperbole is particularly interesting as according to relevance theorists, hyperbole is the most closely related trope to metaphor. They argue that hyperbole differs from metaphor only in the degree to which the broadened interpretation differs from the encoded meaning. In this paper, we report results from a reading-time experiment designed to investigate the processing of metaphor and hyperbole. The main goal of our study was to provide an indication of whether there are processing differences between metaphor and hyperbole.

To preface the experimental part of the paper, we present in more detail, the concept construction accounts of figurative language and discuss their empirical predictions (section 2). Following this, we review some previous experimental studies on metaphor (section 3).

2 Concept construction accounts

Traditionally, metaphorical interpretations have been considered more complex than literal interpretations. A metaphorical interpretation was considered to be derived from the literal interpretation of a sentence (Grice 1989, Ortony 1979, Gentner and Wolff 1997). Empirical studies over the last few decades (see, e.g., Inhoff, Lima and Carroll 1984, Blasko and Connine 1993, Ortony, Schallert, Reynolds and Antos 1978, Blasko and Connine 1993,

Glucksberg, Gildea & Bookin 1982, McElree and Nordlie 1999) have largely refuted this approach as their findings indicated that people do not take any longer to read metaphoric statements than they do to read literal statements**,** which one would have expected if the literal interpretation must be derived before the metaphoric interpretation can be inferred; metaphoric interpretations appeared to be available right from the beginning of processing.

Proponents of truth-conditional pragmatics, such as relevance theorists (Carston 2002, Wilson 2003, Wilson and Carston 2007, Sperber and Wilson 2006) and Recanati (1995, 2001, 2004), have proposed unified ad hoc concept construction accounts of all types of figurative language[1] which fall in line with the empirical research mentioned above, indicating that we should reject the assumption of literalness. Both RTs and Recanati claim that metaphor and approximation require just the same interpretive mechanisms as are used to interpret literal language. These accounts differ both from traditional Gricean analysis (Grice 1989, Ortony 1979, Gentner and Wolff 1997) and from an earlier relevance theoretic account of narrowing and broadening (Sperber and Wilson 1986), where broadened interpretations were considered to be a set of 'implicatures' not explicitly communicated by the speaker.

Both RT and Recanati's concept construction accounts imply that narrowed interpretations and broadened interpretations (metaphors, hyperboles, and approximations) are the opposite outcomes of the same interpretive process of ad hoc concept creation. According to them, when words are used in a less than literal way, the concept communicated by a word is either broader and/or narrower than the concept encoded by that word. A hearer constructs an on-line ad hoc concept out of the salient pieces of encyclopaedic information from the encoded concept. The derived ad hoc concept has either a broader or a narrower denotation than that of the encoded concept. According to the RT account, in the case of broadening, such as hyperbole or metaphor, some content constitutive feature of the lexical concept gets dropped[2], and is no longer a logical property of the ad hoc concept. For instance, in order to interpret the utterance "My bath is boiling" hyperbolically, the content constitutive property *being 100c* in the lexical concept, will not be content constitutive in the ad hoc concept, which means that BOILING*[3] includes in its denotation, baths that are hot but substantially below 100c. Or in order to interpret the metaphor "Call centres are this generation's factories", the property of factories being 'a place of mass production' would be dropped. Whether the resulting interpretation of an utterance is approximate, hyperbolic, or metaphoric, depends on the particular subset of encyclopaedic information that the hearer perceives to be relevant, and therefore uses to construct the on-line ad hoc concept. This is an appealing outcome as it means that figurative language can be discussed as one phenomenon: there are not separate theories of approximation, narrowing, metaphor and hyperbole. Both the RT account and Recanati's account leave open for interpretation what the process of 'dropping' would involve in lexical processing terms. Rubio-Fernandez (2004, 2007) carried

[1] Glucksberg and Keysar (1990) developed a model of metaphor comprehension, which was also based on the assumption that metaphor comprehension involved concept construction. Their model states that interpreting a metaphor involves enhancing attributes of the vehicle that are appropriate for the topic, while suppressing those attributes that are inappropriate. According to this model, the metaphor vehicle stands for a superordinate category rather than referring to the basic-category level. This 'diagnostic category' would include both the metaphoric and the literal referent of the vehicle

[2] Recanati (2001, 2004) argues that concepts have conditions of application packed into the concept rather than properties or features, and it is these conditions which can be manipulated in order to use words in a less than literal sense.

[3] I will be using standard Relevance Theoretic notation in which capital letters are used to mark concepts and * after it is used to mark the fact that it is an ad hoc concept.

out a series of lexical priming studies from which she concluded that, at least in the kind of novel metaphoric materials used in her studies, this dropping process in lexical concept broadening is the result of active suppression (see below for discussion).

As mentioned above, according to both Recanati's and the RT concept construction accounts, both narrowing and broadening contribute to the proposition expressed and therefore affect the truth conditions of the utterance. The ad hoc concept is part of the explicit content of the utterance. Returning to the example "Call centres are this generation's factories", the explicit content of this utterance is "Call centres are this generation's FACTORIES*". The literal interpretation, thus, only has priority over the figurative one, at the local level of words, not at the global level of the utterance. Even then, Recanati and Relevance theorists do not go as far as to say that the literal meaning of the lexical item is fully activated and rejected before the figurative meaning is inferred.

Both accounts propose a parallel processing model of interpretation rather than deriving the figurative interpretation from the literally encoded meaning. RT proposes that the lexically encoded concept is merely accessed in order to search through the encyclopaedic properties to construct the new ad hoc, context specific concept. They propose that the explicit content and the cognitive effects (implicatures) are 'mutually adjusted' in parallel.

> *Mutual adjustment is seen as taking place in parallel rather than in sequence. The hearer does not first identify the proposition expressed then access an appropriate set of contextual assumptions and then derive a set of cognitive effects. In many cases he is just as likely to reason backwards from an unexpected cognitive effect to the context and content that would warrant it.* (Wilson 2003: 34).

Recanati (2001, 2004) argues that the literal concept activated by the linguistic expression, as well as some of the other representations which are activated by association, become possible candidates for the concept that will be selected for interpretation. Although these associated candidates are generated via the literal concept, they are all processed in parallel and compete for activation. The representation that is most active or accessible when the interpretation process stabilizes will be selected, put together and processed with the other components of the utterance, while all other candidates for the meaning of the word are suppressed.

The above aspects of the concept construction accounts have processing implications which we aimed to test by running the experiment reported in this paper. While being sensitive to experimental findings neither RT, nor Recanati, provide clear and precise predictions about processing. There is, therefore, a gap to be bridged between their thesis and empirical studies, such as ours, investigating the cognitive processes involved in utterance interpretation. By interpreting theoretical accounts of language comprehension in processing terms, and designing experiments which can test these predictions, we can glean more information about the underlying cognitive processes involved, which can feed back as premises into theoretical accounts.

The above accounts present several theoretical issues, which are perhaps not directly testable using a reading time methodology. However, finding a significant difference between the reading times of metaphors, hyperboles, and literal statements in a study of the kind we will be reporting in this paper, would give us a preliminary indication that the following aspects of the theories are potentially inaccurate.

- Both accounts propose that the literal interpretation only has priority over the figurative one at the local level of words, not at the global level of the utterance. RT proposes that the 'mutual adjustment of the explicit content and implicatures' takes

place in parallel rather than in sequence. Recanati argues that the associatively related representations and the literal interpretation are all processed in parallel and compete for activation. They are not processed sequentially. Therefore there is nothing in this aspect of either theory which would indicate that figurative interpretations should take any longer to process than literal interpretations. Furthermore, both theories predict that the more implicatures an interpretation generates, the heavier the dependence on memory will be, since the context (previous conversation and surrounding linguistic material) must be recalled in order for the implicatures to be integrated into the discourse. Therefore, the more implicatures there are, the longer the interpretation process will take. However, according to both accounts, metaphoric and hyperbolic interpretations do not necessarily generate any more implicatures than literal interpretations, since the figurative meaning is part of the explicit content.

- Both accounts state that when interpreting hyperbolic and metaphoric utterances, we follow the same comprehension procedure, which involves the online construction of an ad hoc concept. At the level of the theory, no distinction is made between deriving a metaphoric or hyperbolic interpretation. Therefore, other things being equal one would not expect any time difference between the reading of metaphoric and hyperbolic utterances.

The hyperbole and metaphor reading time data which we will be reporting in this paper will allow us to directly contrast the reading times of hyperbole and metaphor in order to see if the above predictions are accurate.

3 Previous experimental research on metaphor

Reading-time studies contrasting the comprehension of literal and metaphoric interpretations (e.g., Ortony, Schallert, Reynolds and Antos 1978, Inhoff, Lima and Carroll 1984), found that provided there is a long enough context, metaphors are processed as quickly as literal statements. Furthermore, sentence verification studies (e.g. Glucksberg, Gildea and Bookin 1982) found that participants were significantly slower at judging literally false sentences as such if there was a true metaphoric interpretation available. The authors concluded that non-literal interpretations are neither optional nor secondary to literal interpretations. Metaphoric interpretations are available right from the beginning of processing; we process literal and non-literal meanings of sentences at the same time. These findings indicate that the Gricean ‘presumption of literalness’ must be flawed. If the literal interpretation of an utterance must be arrived at before the figurative meaning can be inferred, then metaphors should take longer to read than literal statements and the metaphoric interpretations should not interfere with the interpretation and verification of the literal meaning.

Both ad hoc concept construction accounts discussed in this paper, fall in line with these finding. On the basis of their parallel processing models, in which the lexically encoded information and the context are mutually adjusted, there is no reason that metaphors should take longer to read than literal statements. However, both Ortony et al. (1978) and Inhoff, Lima and Carroll (1984), have been criticised as the respective metaphors they used were highly conventionalised, if not idiomatic. Studies have revealed that figurative meanings can become conventionalised, and are then processed more like literal language (Bowdle and Gentner 2005). This draws into question the validity of their results.

Lemaire and Bianco's (2003) findings were the opposite of the two reading-time studies discussed above. They looked at contextual effects on metaphor reading time, and from this they concluded that referential metaphors take longer to process than their literal counterparts but that an inductive context can shorten processing time of referential metaphors. However, there is a potential flaw in their design, as well. Reading time was measured by looking at the difference between 'theoretical reading time' and 'effective reading time' of the reference line. They arrived at the 'theoretical reading time' by calculating the mean reading time of sentences of the same number of syllables. This method is potentially flawed as it doesn't control for word length and word frequency, which can have an effect on reading times (Monsell 1991, Kliegl et al 2004).

There appear to be potential flaws in either the experimental design, or the materials used in the reading time studies discussed above. Therefore, we were interested to see whether the results of our study, would show a reading time difference between literal sentences and metaphors, as we ensured that all metaphor materials were completely novel. In addition, our experiment was designed such that we did not have to calculate 'theoretical reading times'. Participants' reading times of target sentences were straightforwardly recorded.

Other sentence verification studies provide strong support for the concept construction accounts of metaphor. Gernsbacher et al. (1995, 1997, 2001) and McGlone & Manfredi's (2001) findings were the first to indicate that the process of property[4] demotion or dropping described in the concept construction accounts (Carston 2002, Wilson 2003, Wilson and Carston 2007, Recanati 2001, 2004) may be the result of suppression. They found that after reading a metaphorical statement, subjects were faster to verify statements related to the metaphoric interpretation of the vehicle, and slower to verify statements related to the literal interpretation of the vehicle, when compared to the time it took them to respond to the same statements after reading a control sentence. They therefore concluded that metaphor does involve enhancing attributes that are relevant to the metaphorical interpretation of the vehicle while suppressing those that are irrelevant, as is proposed in the concept construction accounts of figurative language (Carston 2002, Wilson and Carston 2007, Recanati 2004).

Rubio-Fernandez (2004, 2005, 2007) used a cross-modal priming paradigm, rather than the sentence verification task used in the Gernsbacher studies. This made it possible to get an on-line measure of property activation across time, and to determine at which point in processing, there is a loss of activation of irrelevant literal properties of the metaphor vehicle. Her findings also suggest that the 'dropping' process described in the concept construction accounts of broadening are the result of suppression. Swinney (1979) found that the irrelevant meanings of homonyms were suppressed as early as 200-300ms from the offset of the ambiguous words. Rubio-Fernandez (2004, 2005, 2007) however, found that in metaphor conditions, there was loss of activation of superordinates between 400ms and 1000 ms as a result of active suppression of metaphor irrelevant properties. She concluded from these results that suppression of metaphor inconsistent features takes place during metaphor processing, but that although suppression is involved in both disambiguation and metaphor interpretation, it acts differently in the two different processes because the former is a case of meaning selection, while the latter is a case of meaning construction. Rubio-Fernandez (2004, 2007, 2008) argues that the difference in time course of suppression of irrelevant properties or meanings in disambiguation and metaphor interpretation is due to the different pragmatic processes involved. She argues that the attentional processes involved in novel metaphor interpretation involve some level of awareness and need time to develop; whereas the

[4] Or 'Conditions of application' for Recanati (2001, 2004)

suppression involved in disambiguation does not require this level of attention and awareness. Therefore suppression operates faster in disambiguation than in metaphor interpretation. This could be an important factor to consider when contrasting metaphor with other tropes such as hyperbole. We know from Rubio-Fernandez's studies that metaphor comprehension involves this attentional suppression because metaphor irrelevant features are frequently not just irrelevant but also inconsistent and in direct conflict with the metaphoric meaning. This raises the question of whether active attentional suppression is required in the processing of other tropes such as hyperbole, as hyperbole irrelevant features of the lexical concept may not be inconsistent with the hyperbolic meaning, but perhaps just irrelevant. If hyperbole does not require active attentional suppression as metaphor does then we may find that metaphors take longer to read than hyperboles.

There is little, if any, research comparing metaphor and hyperbole. However, a small study we carried out as part of previous work (Unpublished MRes- Deamer 2009) suggests that there are processing differences between the two degrees of broadening; approximation and metaphor. The study explored the predictions of the unified account of figurative language by testing 5 year old children's performance at interpreting novel metaphors, conventionalised metaphors, approximations and narrowings and contrasting their performance on each of the phenomena. Children performed significantly better at approximation than they did at narrowing, conventionalised metaphor or novel metaphor. They also performed significantly better with conventionalised metaphor than novel metaphor. These results make sense if these metaphors require a process of disambiguation instead of ad hoc concept construction. The same group of children were anecdotally recorded as having already begun to produce novel hyperboles in their everyday speech. These results are interesting because they show a disparity in young children's performance on approximation and novel metaphor, which Carston and her colleagues (Carston 2002, Wilson and Carston 2007), as well as Recanati (2004), propose to be interpreted via the same comprehension procedure. The children's differing performances on the metaphors and approximations does not give us any clear evidence that ad hoc concept construction is not required in order to interpret all three tropes, but it does suggest that there might be processing differences between these three tropes, and that these differences cause 5 year old children to struggle with metaphor, but not with approximation. These off-line results and the anecdotal evidence provide motivation to further investigate the processes involved in the comprehension of these different degrees of broadening using an on-line methodology.

The concept construction theories propose to account for hyperbole (among other tropes) as well as metaphor, but previous on-line studies (Rubio-Fernandez 2004, 2007, 2008, Gernsbacher et al. 2001, McGlone and Manfredi 2001) have only focused on metaphor. Thus, questions regarding how well the concept construction accounts accurately capture the processes and mechanisms underlying hyperbole comprehension remain unanswered. Our endeavour in the present study is to contrast the processing of metaphor and hyperbole. The on-line cross modal priming paradigm Rubio-Fernandez (2005, 2007, 2008) used was perfectly suited to test for property activation during metaphor comprehension. However, this paradigm is very difficult to adapt to test property activation during hyperbole comprehension. Rubio used superordinates of the target word, as metaphor inconsistent properties (e.g. Cheetah- CAT). Additionally, she used distinctive properties of the target word as the metaphor consistent properties. Unfortunately, many hyperbolic terms do not have superordinates, and often they do not have distinctive properties either (e.g., He couldn't run anymore his knee was **smashed**). In light of this, we chose to use a reading time paradigm as

a first attempt at comparing hyperbole and metaphor comprehension, and explore whether both involve the same process of ad hoc concept construction.

4 Reading time experiment

We carried out a reading-time experiment, comparing the reading times of target sentences when preceded by, either, a metaphoric, a hyperbolic, or a literal biasing context. Each target sentence contained a specific target word (e.g. 'it was an **assault'**) which was to be interpreted literally, hyperbolically or metaphorically depending on the biasing context. We also compared the reading times of the different tropes preceded by either a short context or a long context length. Indeed, we thought that context length may affect reading times: a longer context provides more time for appropriate features to be activated, and therefore perhaps lessens the need for suppression of irrelevant features, in the case of metaphor and hyperbole. Ortony et al. (1978) found that metaphors took no longer to read than literal statements, provided that the context was long enough. In the short context condition, metaphors took significantly longer to read than literal statements. The authors' interpretation of this interaction was that the short context condition did not provide appropriate schemata for interpreting the metaphoric targets.
The unified concept construction accounts of figurative language discussed above (Carston 2002, Wilson and Carston 2007, Recanati 1995, 2001, 2004) would not predict there to be any difference in reading times between metaphors and hyperboles, other things being equal. Furthermore, previous studies have shown that there is no such reading time difference between metaphoric and literal utterances. It was this set of predictions which we aimed to test in this study

4.1 Method
4.1.1 Participants

Sixty native English speaking adults took part in this study. They were recruited through the University College London psychology subject pool, and received a £3 compensation for their time.

4.1.2 Procedure

We measured the time it took participants to read a target sentence under different conditions of target interpretation (in hyperbolic, metaphoric, literal and unrelated contexts). We used a now standard reading time methodology (see, e.g., Ortony et al. 1978 and Inhoff et al. 1984 among many others) in which participants were asked to read a series of short target sentences preceded by a literal, hyperbolic or metaphoric context, on a laptop screen. The target sentences contained a target word which should have been interpreted either literally, hyperbolically or metaphorically, depending on the biasing context (e.g. 'It was an **assault**'). In the present study, the preceding contexts were short (5-13 words) for half the participants, and long (55-86 words) for the other half of the participants. After reading and understanding the context passages, which appeared in one complete passage on a single computer screen, participants were asked to press the spacebar to bring up a new screen on which the target sentence appeared. Participants were asked to press the spacebar again when they had read and understood the sentence. Following one third of the target sentences, they were then asked to answer a comprehension question (e.g., for metaphor context and target sentence 'our new flat had a huge silver, double fridge- It was a monster', the comprehension question was 'do you think the fridge is alive'). This was to ensure that they were focused on

understanding the passages and not just reading them. Target sentences were randomly selected to be followed by a comprehension question. Before the critical materials began, participants saw five practice trials. These trials all had the same structure as the critical trials, and they all had literal interpretations.

Participants were tested individually. Prior to the start of the experiment, each participant received instructions regarding the procedure (i.e. what buttons to press and what they will see on the screen). They were told that the experiment is a simple, computer-based experiment which is looking at how we process language. No information about the exact relationship between the context and target was provided.

4.1.3 Materials

All stimulus materials were designed for the purposes of this experiment. For half of the participants, the target words were preceded by short contexts, for the other half they were preceded by long contexts. See Table 1 below for some examples of contexts used.

Table 1: Example of short and long context material

Short Literal Context	**Target sentence**
Sam and Mark went for a walk in a national park.	*It was a forest.*
Long Literal Context	
Sarah and Mark loved doing things together at the weekend. Sometimes they went shopping. Sometimes they decided to go and visit friends in different cities. They didn't get to spend that much time with each other during the week so they looked forward to the weekends. This weekend it was sunny and they went for a walk in the local national park.	
Short Hyperbolic context	
The back yard definitely needed pruning.	
Long Hyperbolic Context	
When they had bought their house they had been attracted to it because of the size of the garden. They loved how it was big enough for the children to kick a ball around and have some fun. Their previous garden had been very small. They had not bargained for how much work it would involve though. It needed pruning all the time.	
Short Metaphoric context	
Sam always got lost. The university was enormous.	
Long Metaphoric context	
Sarah had begun her first year at university. Even though she had gone to a large state school, nothing could have prepared her for the size of her university campus. It took her a while to even find where her department was. In her first weeks, she was always getting lost. The university was enormous and pretty intimidating.	

Short contexts ranged from 5-13 words in length, whereas long contexts ranged from 55-86 words. The target sentences were designed using the formal criteria below;
(a) In all of the descriptions (contexts and target sentences combined), an object or concept in the context, is described using the target word in the target sentence (e.g. A fridge is described as a monster, or earl grey tea is described as noxious).
(b) The hyperbolic and metaphoric descriptions differed in one way: A metaphorical description can never be literally true no matter how much you manipulate the context. However, a hyperbolic description can be literally true if the context is modified slightly. In the hyperbolic descriptions, if the contexts were changed slightly, the object or concept depicted in the context, could be literally described using the target word in the target sentence (e.g. earl grey tea could be described as noxious if the context indicated that it had been spiked with a drug). However, in the metaphoric descriptions, the object or concept depicted in the context, could never be literally described using that target word, even if the context were changed in some way to facilitate this (e.g. a fridge can never be a monster).

The contexts were all written with the following considerations in mind;

a) The metaphor, hyperbole and literal contexts clearly induce either a metaphorical, hyperbolic or a literal interpretation.

b) The contexts themselves were written using only literal language. Any loose uses which may be included were very conventionalised.

c) The contexts did not just translate the target sentence into literal language.

d) All efforts were made to ensure that the degree to which the target sentence followed the context was as similar as possible in the metaphoric, hyperbolic and literal conditions.

How well the contexts we constructed adhered to the guidelines (a-d) above was assessed through a questionnaire. We presented 27 target sentences to 38 native English speaking participants, recruited via email. Each participant saw 7 of the target sentences in metaphoric contexts, 7 in hyperbolic, and 7 in literal contexts. Twenty-one participants saw the target sentences in short contexts and 16 saw them in long contexts. They were asked to decide whether a certain target sentence should be interpreted literally, hyperbolically or metaphorically. They were then asked to rate how much the target sentence related to the context on a scale of 1- 7 (1 being if the context and target sentence bear no relation to one another, and 7 being if they are extremely related). If a target sentence was consistently misinterpreted, then it was rejected and removed from the experimental materials. If it was fairly consistently correctly interpreted in two context conditions, but was often misinterpreted in one of the context conditions, then we changed that context, in order that it more clearly biased the intended interpretation[5]. These changes were checked with a group of eight UCL linguists, who all verified the materials. Any contexts which were biasing a literal

[5] One problem which was detected after the questionnaire was run, was that the first task relied upon people having a clear understanding of the distinction between hyperbole and metaphor, which they did not. Although it was explained what hyperbole was at beginning of the questionnaire, it is a complex notion to comprehend and use accurately in a task of this nature. This invalidated their responses on whether they interpreted something metaphorically or hyperbolically. When participants stated that they interpreted a target sentence in a metaphor biasing context, hyperbolically, or vice versa, this could not be taken as an indication of whether this was in fact the case, because they could not reliably distinguish between hyperbole and metaphor. It was however, important to get the data from the questionnaire, to ensure that people were at least correctly interpreting the target sentences which were in literal biasing contexts, literally, and that they were not interpreting either hyperbolic or metaphoric sentences, literally. We did not always change or reject target sentences in which participants had interpreted them metaphorically in hyperbolic biasing contexts, or vice versa, because of the issue discussed above.

interpretation, in a figurative condition, or a figurative interpretation, in a literal condition, three or more times, were changed, and the target item was kept. If they were miss-interpreted in this way more than five times, then the whole target item was removed. The target items were also removed if it was impossible to change the context which was inducing miss-interpretations. All target items were checked by the same group of scholars, and a whole target item was also removed if the feedback from the group had been that the item was too conventional, or did not adhere to the formal criteria of the distinction between hyperbole and metaphor (discussed above). Additionally, from the questionnaire responses to the sliding scale question, we calculated the mean sliding scale score (between 1 and 7) for each target item in each context. If a target item had a mean sliding scale score of below 3 in any of the contexts then it would not have been included. All target items had a mean sliding scale of above 3, and were therefore not rejected on that basis.

We also ensured that all contexts were free of any words which may have had a priming effect on the target sentence. In order to achieve this, a questionnaire based on the research of Rosch and Mervis (1975) and Barsalou (1987), was used to identify strong associates of the target words within the target sentences, which may have had a priming effect. The questionnaire comprised two different tasks: a brief definition task and a free association task. The questionnaire included a list of 27 target words which could potentially have been included in the critical target sentences in the experiment (i.e. *chewed, noxious, novel, tree,* etc.), of which 21 were finally selected for the study. For each target word, participants were asked to first, briefly define the word, and then secondly, list up to 5 words which they associated with the target word. The questionnaire was run on 21 native English speaking adults also recruited via email. None of the words which were present in the participants' answers were included in the contexts. This ensured that no words which participants either associated with the target words, or used to define the target words, were included in the contexts.

4.1.4 Counterbalancing

Three lists were created, each list containing the same 21 target sentences. The three lists differed with respect to their prior contexts. In the first list, 7 target sentences were preceded by a context consistent with the literal interpretation, 7 were preceded by metaphoric consistent contexts, and 7 were preceded by hyperbolic consistent contexts. In the second and third lists, the target sentences appeared in different contexts, so the target sentences which were placed in a literal context in list 1, were placed in a context consistent with a metaphoric interpretation in the second list, and in the third with a hyperbolic interpretation. If a target sentence (e.g. *It was noxious)* was preceded by a metaphoric consistent context (*The politicians style of politics was typical)* in the first list, then it was placed in a literal context (*The varnish he was using was a new brand)* in the second list, and a hyperbolic context (*He had to drink earl grey and he hated it)* in the third list, and so on. Each list contained seven passages in each of the three conditions (literal, metaphoric, hyperbolic). The order in which the target sentences appeared was randomized for each participant. The trope condition was within subjects, so each participant was randomly assigned to one list so that they only saw each target sentence once, in one condition. The context length condition was between subjects, so half of the participants saw the target sentences in long contexts, and half saw them in short contexts.

4.1.5 Apparatus

The experiment was carried out using a small net book computer, positioned on a table. Participants sat on a chair, up to the table and operated the computer. The texts were presented in lower case, font size 24. The experiment was programmed using e-prime.

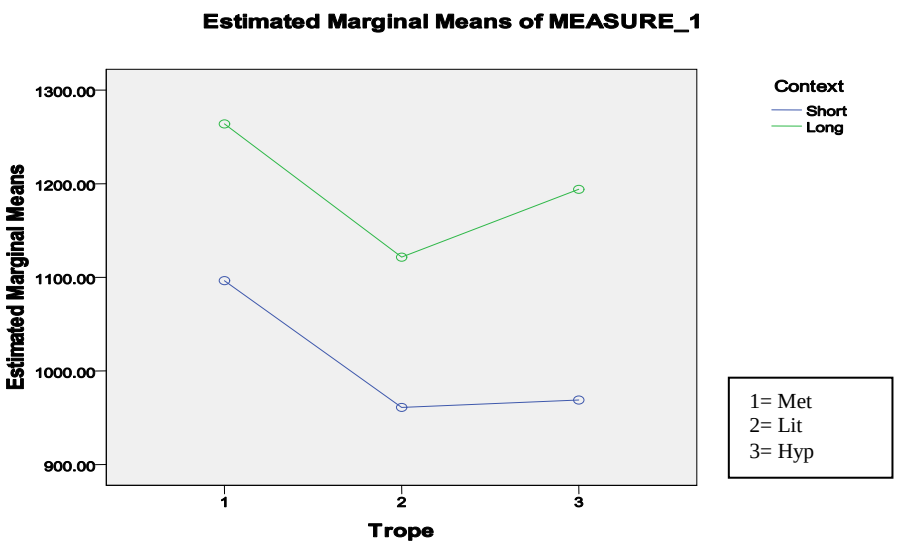
4.1.6 Data analysis

The dependent measure in this experiment was sentence reading time. The independent variables were type of trope, and context length. Trope was within-subjects, and context length was between subjects. The sentence reading time was calculated by the length of time from the onset of the target sentence text, to the response on the key board.

4.1.7 Results and analysis

Participants reading times of target sentences in the three different trope conditions, and the two context length conditions are illustrated in Figure 1.

Figure 1. Mean reading times



Analysis of variance in reading times, between (context length condition) and within (trope conditions) groups, was analysed using a split plot ANOVA (with participants as the random variable (F_1) and items as the random variable (F_2)), with context length group as the between subjects factor, and trope as the within subjects factor. A main effect of the context length ($F_1(1, 56)=5.69$ $p=.020$, $F_2(1, 40)=20.49$ $p<.001$) indicated a difference between reading times in the long and short conditions. We also found a main effect of trope ($F_1(2, 112)=15.57$ $p<.001$, $F_2(1.89, 75.77) = 9.54$, $p<.001$) pointing to a difference between reading times in the three trope conditions. Yet, there was no significant interaction between factors ($F_1(2, 112)=.959$, $p=0.387$, $F_2(1.89, 75.77)=1.49$, $p=.231$), indicating that the relationships between all three tropes were the same in both long and short contexts.

As illustrated in Figure 1, participants took significantly longer to read the target sentences in the long condition than in the short condition. The most obvious interpretation of this result is that the longer contexts required a heavier dependence on memory and generally

more processing effort. In the long contexts, the participants had more context to recall in order to mutually adjust the context and the explicit content of the utterance in order to arrive at its final interpretation. Furthermore, the extra information contained in the longer contexts would mean that there were likely to be more inferences drawn. This will have caused the reading times to be longer in the hyperbolic and literal conditions, as we can observe.

Post hoc Bonferroni corrected pairwise comparisons indicate that target sentences took longer to read when preceded by a metaphoric biasing context than when preceded by a literal biasing context ($p<.001$). This finding could be explained by the role of suppression in metaphor processing. In order to arrive at a metaphoric interpretation, metaphoric inconsistent features of the encoded concept must be actively suppressed. Rubio- Fernandez's (2004,2007) data provide strong evidence for the presence of the kind of suppression of irrelevant features, which both the RT account and Recanati's account propose to be a key part of figurative language interpretation. Her data also indicate that the suppression involved in metaphor interpretation is a more attentional process which occurs later on in processing than the automatic suppression involved other pragmatic processes such as disambiguation, and could therefore cause the interpretation of a metaphoric utterance to take longer than that of a literal utterance, as our data indicate. The lack of contextual information in the short contexts cannot account for the observed reading time difference, as suggested by previous studies (Ortony et al 1978, Inhoff et al. 1984) which indicated that metaphor interpretation benefits from a longer context, because we found no interaction between trope and context length; the difference observed between metaphorical and literal reading times were similar in long and short contexts.

Post hoc pairwise tests also revealed significantly longer reading times for metaphors over hyperboles (Bonferroni, $p=.002$). Unlike what RT and Recanati's thesis leads us to expect, this indicates processing differences between hyperbole and metaphor. One way of explaining these differences is to propose that the suppression of inconsistent core features plays less of a role in hyperbole comprehension than in novel metaphor comprehension. If, as Rubio-Fernandez (2004, 2007) argues, suppression of metaphor inconsistent core features is an attentional process which takes longer to develop than other pragmatic processes, and this type of suppression is only required in metaphor comprehension, then this would explain why hyperboles are quicker to read than metaphors.

Finally, further post hoc pairwise tests revealed no significant difference between reading times of hyperboles and literal sentences (Bonferroni, $p=.245$). This further supports the argument that hyperbole comprehension does not require the effortful and attentional process of suppression. Hyperboles appear to be processed more like literal sentences than metaphors.

5 General Discussion

Participants in our study read target sentences that could be interpreted metaphorically, literally or hyperbolically, depending on the bias of the preceding context. Overall, we found that when the context was longer, response times on target sentences were longer. This result is likely to be due to the fact that richer contexts require a greater call on memory when deriving inferences, and are likely to yield more inferences. Post-hoc analyses for the significant main effect of trope revealed that, when biased to be interpreted metaphorically, sentences took longer to read than in the literal or hyperbolic condition, but there was no difference between the literal and hyperbolic conditions.

The observed difference between the reading times of the target sentences in the literal and metaphor conditions seems to be in contradiction with earlier findings (Ortony et al., 1978) and Inhoff et al., 1984). They found that metaphors took no longer to read than literal statements, provided that the context was such that it heavily biased the metaphoric

interpretation, and provided that the context was long enough. One possible explanation for our differing findings, is that the metaphors used in the above studies were conventionalized, if not idiomatic (e.g. "regardless of the danger, the troops marched on"), whereas the metaphors and hyperboles designed for our study were completely novel. Bowdle and Gentner (2005, among others) show that figurative meanings can become conventionalised, and are then processed more like literal language. Such conventionalised metaphors are considered to be instances of polysemy, in which both the figurative and literal readings have been stored in the lexicon as separate entries. If the metaphors used in Ortony et al.'s study were conventionalised, then the reading times for metaphors may not have differed from literal statements, because the metaphors were in fact being processed as literal statements.

As mentioned in Section 2, it is not obvious why metaphoric interpretations would necessarily carry more implications than literal interpretations. Thus, this cannot explain why metaphors take longer to read than literal sentences. If we assume with these new ad hoc concept theories that metaphor contributes to the proposition expressed, then the observed difference between reading times in the metaphor condition, and the literal condition, cannot be attributed to the heavier dependence on memory, and the longer mutual adjustment procedure. We must, then, assume that the cause of the reading time differences in the metaphor and literal conditions is located in the processes involved in deriving the proposition expressed.

In contrast to the difference between the metaphoric and literal conditions, our results do not show a significant difference between the reading times of hyperboles and literal sentences, thus indicating that there is something unique about the processing of metaphor, which caused metaphors to take longer to interpret than hyperboles and literal sentences. Rubio- Fernandez's (2004, 2007) findings provide strong evidence for the presence of the kind of suppression of irrelevant features in metaphor comprehension, which explains in lexical processing terms, the 'dropping process' which both the RT account and Recanati's account propose to be a key part of figurative language interpretation. Rubio-Fernandez concluded from her findings that the kind of suppression involved in metaphor interpretation is a more attentional process which occurs later on in processing than the automatic suppression involved in other pragmatic processes such as disambiguation. It would, therefore, be plausible that the interpretation of a metaphoric utterance would take longer than that of a literal utterance, which corresponds to what we find.

Our results provide evidence that metaphors take longer to read than literal sentences, because they require the active suppression of metaphor irrelevant features. In addition, the lack of difference between the reading times of hyperboles and literal sentences, and the significant difference between the reading times of hyperboles and metaphors, suggest that hyperbolic interpretations do not draw on costly attentional suppression processes. This is a tentative conclusion, as despite our efforts to control for other factors which could affect reading times, such as degree of context bias, it is possible that the metaphor biasing contexts were such that they induced richer interpretations, which could account for the reading time difference.

If metaphors and hyperboles do require a different style of processing, then this is not predicted by the concept construction accounts. Our finding is not inconsistent with the broad outline of either the RT or Recanati's account, as both hyperbolic and metaphoric interpretations could both still be the result of the construction of an occasion specific concept. However, our finding is inconsistent with a more detailed account, according to which the process of 'dropping' or 'demotion' of inconsistent features in both metaphor and hyperbole, is the result of active suppression.

This interpretation of our data needs to be further investigated in a follow up study which will enable us to specifically look at the activation levels of figurative inconsistent

features, at different stages of processing of both hyperboles and metaphors. We plan to carry out a sentence-verification task similar to that carried out in Gernsbacher et al. (2001) in which participants will be asked to read metaphor, hyperbole and literal prime sentences, before verifying a target property statement which is relevant to either the superordinate or the basic level category of the target word in the prime sentence. By comparing verification latencies, we will be able to observe suppression effects in the different conditions

References

Barsalou, L. 1987. The instability of graded structure in concepts. In U. Neisser (Ed.)
Concepts and conceptual development: Ecological and intellectual factors in categorization 101-140). New York: CUP.

Blasko, D. G., and Connine, C.M.1993. Effects of familiarity and aptness on metaphor
processing. Journal of experimental psychology: Learning, Memory, and Cognition,19,295-308.

Bowdle. B, and Gentner. D. 2005. The Career of Metaphor. Psychological Review. Vol
112(1), Jan 2005, 193-216

Cacciari and Glucksberg. 1994. Understanding figurative language. In M. Gernsbacher (ed.) Handbook of psycholinguistics. Academic press. (pp. 447-77).

Carston, R.2002. Thoughts and utterances. Chapter 5. Blackwell publishing

Fodor, J 1998. Concepts: where cognitive science went wrong-Oxford university press

Gentner, D., & Wolff, P. 1997. Alignment in the processing of metaphor. Journal of memory and language, 37, 331-355.

Gernsbacher, M.A., Keysar, B., and Robertson, R. R. 1995. The role of suppression in metaphor interpretation. Paper presented at meetings of the psychonomic society, LA, CA

Gernsbacher. M.A, Keysar. B, Robertson. R, and Werner. N 2001. The role of suppression and enhancement in understanding metaphors. Journal of Memory and Language. 45- 433-450

Gibbs, R.1994. Poetics of Mind- Figurative thought, language and understanding.CUP

Grice, P.1989. Studies in the ways of words. Harvard University Press.

Glucksberg, S.2001. Understanding figurative language: from metaphors to Idioms.OUP

Glucksberg, S.2003. The psycholinguistics of metaphor. Trends in cognitive sciences. 7:92-6

Glucksberg, S, P, Gildea & H.B Bookin. 1982. On understanding non-literal speech: Can people ignore metaphors? Journal of verbal learning and verbal behaviour. 21:85-98

Glucksberg, S &B. Keysar. 1990. Understanding metaphorical comparisons: Beyond similarity. Psychological review. 97: 3-18

Glucksberg, S., McGlone, M. S., and Manfredi, D. 1997. Property attribution in metaphor comprehension. Journal of Memory and Language, 36, 50-67.

Inhoff, A., Lima, S., and Caroll, P. 1984. Contextual effects on metaphor comprehension in reading. Memory and Cognition, 12,558-567.

Kliegl. R, Grabner. E, Rolfs. R, Engbert. R. 2004. European journal of cognitive psychology. 16 (1/2), 262-284.

Lemaire. B and Bianco. M (2003). Contextual effects on metaphor comprehension: Experiment and simulation. Conference on Cognitive modeling.

McElree, B., & Nordlie, J. 1999. Literal and figurative interpretations are computed in parallel. Psychonomic bulletin and review. 6, 486-494.

McGlone, M.S., Glucksberg, S., and Cacciari, C. 1994. Semantic Productivity and Idiom
comprehension. Discourse Processes, 17, 167-190.

McGlone, M.S., and Manfredi, D. 2001. Topic-vehicle interaction in metaphor comprehension. Journal of Memory and Language. 29 (8), 1209-19.

Monsell, S. 1991. The nature and Locus of word frequency effects in Reading. In Basic processes in reading: Visual word recognition. 1991.

Ortony, A. 1979. Beyond literal similarity. Psychological review 86, 161-180

Ortony, A., Schallert, D., Reynolds, R., and Antos, S. 1978. Interpreting metaphors and idioms: Some effects of context on comprehension. Journal of verbal learning and verbal behavior,17, 465-477.

Recanati, F.1995. The alleged priority of literal interpretation. Cognitive science.19, 207-232

Recanati, F. 2001. What is said. Synthese. 128, 75-91

Recanati, F. 2004. Literal meaning. CUP

Rosch. E, and Mervis. C.B. 1975. Family resemblances: Studies in the internal structure of categories. Cognitive psychology. 7 (4), 573-605

Rubio-Fernandez, P. 2004. Unpublished PhD Thesis. Cambridge University

Rubio-Fernandez, P. 2007. Suppression in metaphor interpretation: Differences between meaning selection and meaning construction. Journal of semantics 24: 345-371

Rubio-Fernandez, P. 2008. Concept narrowing: The role of context independent information. Journal of semantics. 25, 381-409.

Sperber, D & Wilson. 1986, 1995. Relevance: communication and cognition. Blackwell.

Sperber, D & Wilson. 2006. A deflationary account of metaphor. UCL WPL.

Swinney, D. 1979. Lexical access during sentence comprehension(Re) consideration of context effects. Journal of verbal learning and verbal behavior.18; 645-659 (1979)

Tabossi and Zardon. 1991. Processing ambiguous words in context. Journal of memory and language. 32: 359-372

Wilson, D.2003. Relevance and lexical pragmatics. Italian journal of linguistics 15.2:273-291

Wilson, D and Carston, R.2007. A unitary approach to Lexical Pragmatics: Relevance, Inference and Ad hoc concept. In Noel-Roberts (ed.) 2007. Pragmatics. Palgrave, London:230-259.

Wilson, D and Sperber, D. 2002. Truthfulness and Relevance. Mind. OUP.